HHS Public Access

Author manuscript

Pharmacogenet Genomics. Author manuscript; available in PMC 2022 July 01.

Published in final edited form as:

Pharmacogenet Genomics. 2021 July 01; 31(5): 116–123. doi:10.1097/FPC.0000000000000429.

Genome-wide association study of letrozole plasma concentrations identifies non-exonic variants that may affect CYP2A6 metabolic activity

Daniel L. Hertz, PharmD, PhD[1], **Julie A. Douglas, PhD**[2,3], **Kelley M. Kidwell, PhD**[4], **Christina L. Gersch**[5], **Zeruesenay Desta, PhD**[6], **Ana-Maria Storniolo, MD**[6], **Vered Stearns, MD**[7], **Todd C Skaar, PhD**[6], **Daniel F Hayes, MD**[5], **N. Lynn Henry, MD, PhD**[5], **James M. Rae, PhD**[5]

[1]Department of Clinical Pharmacy, University of Michigan College of Pharmacy, Ann Arbor, MI, United States, 48109-1065

[2]Department of Human Genetics, University of Michigan Medical School

[3]Department of Mathematics and Statistics, Skidmore College, Saratoga Springs, NY, United States, 12866

[4]Department of Biostatistics, University of Michigan School of Public Health

[5]Department of Internal Medicine, Division of Hematology/Oncology, University of Michigan Medical School

[6]Indiana University

[7]Johns Hopkins School of Medicine

Abstract

Objective: Letrozole is a non-steroidal aromatase inhibitor (AI) used to treat hormone receptor positive (HR+) breast cancer. Variability in letrozole efficacy and toxicity may be partially attributable to variable systemic drug exposure, which may be influenced by germline variants in the enzymes responsible for letrozole metabolism including Cytochrome P450 2A6 (CYP2A6). The objective of this genome-wide association study (GWAS) was to identify polymorphisms associated with steady state letrozole concentrations.

Methods: The Exemestane and Letrozole Pharmacogenetics (ELPh) Study randomized post-menopausal patients with HR+ non-metastatic breast cancer to letrozole or exemestane treatment. Germline DNA was collected pre-treatment and blood samples were collected after 1 or 3 months

Corresponding Author: Daniel L Hertz, 428 Church St., Room 3054 College of Pharmacy, Ann Arbor, MI 48109-1065, Office phone: (734) 763-0015, Fax: (734) 763-4480, DLHertz@med.umich.edu.

This work was presented in part at the 2020 American Society of Clinical Oncology

Conflict of Interest

This work was supported in part by Pfizer and Novartis Pharma AG. Dr. Stearns has received research funding from Abbvie, Celgene, Merck, Novartis, Medimmune, Pfizer, and Puma. Dr. Henry has received research funding to conduct pharmaceutical sponsored clinical trials from Abbvie, Innocrin Pharmaceuticals, and Pfizer. Dr. Hayes reports research funding from Merrimack Pharmaceuticals, Eli Lilly, Menarini Silicon Biosystems, Puma Biotechnology, Pfizer, and Astra Zeneca in the last 24 months. He also reports consulting fees from Cepheid, Freenome, Artiman Ventures, Agendia, Lexent Bio, Epic Sciences, and Salutogenic Innovations. He is the named investigator of a patent held by the University of Michigan which is licensed to Menarini Silicon Biosystems, from whom he receives annual royalties. He holds stock options in Oncimmune LLC and InBiomotion.

of treatment to measure steady-state letrozole (and exemestane) plasma concentrations via HPLC/MS. Genome-wide genotyping was conducted on the Infinium Global Screening Array (>650,000 variants) followed by imputation. The association of each germline variant with age- and body mass index-adjusted letrozole concentrations was tested in self-reported white patients via linear regression assuming an additive genetic model.

Results: There were 228 patients who met the study specific inclusion criteria and had both DNA and letrozole concentration data for this GWAS. The association for one genotyped polymorphism (rs7937) with letrozole concentration surpassed genome-wide significance ($p=5.26x10^{-10}$), explaining 13% of the variability in untransformed steady-state letrozole concentrations. Imputation around rs7937 and in silico analyses identified rs56113850, a variant in the CYP2A6 intron that may affect CYP2A6 expression and activity. rs7937 was associated with age- and body mass index-adjusted letrozole levels even after adjusting for genotype-predicted CYP2A6 metabolic phenotype ($p=3.86x10^{-10}$).

Conclusions: Our GWAS findings confirm that steady-state letrozole plasma concentrations are partially determined by germline polymorphisms that affect CYP2A6 activity, including variants near rs7937 such as the intronic rs56113850 variant. Further research is needed to confirm whether rs56113850 directly affects CYP2A6 activity and to integrate non-exonic variants into CYP2A6 phenotypic activity prediction systems.

Keywords

Pharmacogenetic; genome-wide association study (GWAS); letrozole; pharmacokinetics

Introduction

Aromatase inhibitors (AI) are a class of agents commonly used in patients with hormone receptor positive (HR+) breast cancer[1,2]. AIs inhibit the aromatase-mediated conversion of androgens to estrogens, depleting systemic estrogen concentrations[3] and depriving HR+ tumors of their estrogenic growth factor. Along with their effectiveness, AI cause toxicities that resemble the effects of estrogenic deprivation during menopause[4]. These toxicities, notably musculoskeletal (i.e., arthralgias and myalgias) and vasomotor (i.e., hot flashes) symptoms, necessitate treatment discontinuation in about a quarter of AI-treated patients[5].

Inter-patient differences in AI tolerability and/or estrogenic response may be due, in part, to differences in circulating AI concentrations during treatment[6,7]. Prior work from our group, and others, have identified clinical and genetic predictors of circulating AI concentrations during treatment[8]. Pharmacogenetics analyses of candidate single nucleotide polymorphisms (SNPs) conducted in the Exemestane and Letrozole Pharmacogenetics (ELPh) study have found that circulating plasma concentrations of exemestane and letrozole are affected by inherited SNPs in *CYP3A4*[9] and *CYP2A6*[10], respectively. Still, only a small proportion of the variability in systemic drug concentration is explained by *CYP3A4* and *CYP2A6* genotype, even after accounting for clinical factors such as age and body mass index (BMI).

Genome-wide association studies (GWAS) can confirm pharmacogenetic associations previously detected in candidate SNP studies or discover novel associations in genes not previously suspected to be associated with the phenotype[11]. For example, a GWAS of circulating concentrations of anastrozole, an AI that is chemically and pharmacologically similar to letrozole, implicated a SNP (rs11648166) located in a previously unsuspected anastrozole influx transporter (SLC38A7)[12]. The objective of this investigation was to conduct a GWAS of patients in the ELPh trial to further assess the association between *CYP2A6* and letrozole levels and investigate whether any other genes, including *SLC38A7*, contribute to inter-patient variability in letrozole concentrations during treatment.

Materials and Methods

ELPh Patients and Treatment

The Consortium on Breast Cancer Pharmacogenomics (COBRA) conducted the prospective, open-label, ELPh study[13], which enrolled post-menopausal women with stage I-III HR+ breast cancer from Indiana University Cancer Center, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, and the University of Michigan Comprehensive Cancer Center from August 2005-July 2009. Eligible patients were considering AI therapy upfront or following tamoxifen after completion of local therapy (i.e., surgery and/or radiation) and systemic chemotherapy. Patients were stratified by prior bisphosphonate, tamoxifen, and chemotherapy treatments and randomized 1:1 to receive oral exemestane (25 mg/day) or letrozole (2.5 mg/day) for 2 years. The Institutional Review Boards of each site approved the protocol and all patients provided written informed consent prior to enrollment.

Circulating Letrozole Concentrations

Circulating letrozole plasma concentrations were measured in samples collected after 3 months of AI treatment, or after 1 month in patients who crossed-over to the alternative treatment arm, as previously reported[10]. Blood samples were collected in heparinized tubes approximately two hours after the patient took their daily AI dose to estimate a steady-state maximum systemic concentration (C_{max}). Letrozole plasma concentration was measured via high performance liquid chromatography (LC) with fluorescent detection with LLOQ = 7.0 ng/ml, as previously described[10].

Genome-wide Genotyping and Imputation

Germline DNA was isolated using the QIAamp DNA Blood Maxi Kit–Spin (Qiagen, Valencia, CA) from a whole blood sample collected at enrollment[10]. Germline DNA was sent to the University of Michigan Advanced Genomics Biomedical Research Core for genome-wide genotyping on the Infinium Global Screening Array, which contains more than 650,000 variants, including a genome-wide backbone (>530,000 variants) and curated clinical variants, most notably pharmacogenomic candidate SNPs. Given the small number of non-white patients enrolled in the ELPh trial, only self-reported white patients were included in this GWAS. Genotype quality control was conducted to eliminate variants with low call rates (<95%), monomorphic variants, and variants for which the observed genotype distribution departed from Hardy Weinberg Equilibrium ($p<10^{-6}$). Sample call rates ranged from 98.24 to 99.95%. More than 16 million variants were also imputed from the Haplotype

Reference Consortium (HRC) panel (r1.1.2016) using the Michigan Imputation Server.[14] Imputation was performed using the European reference panel (1000G Phase3 EUR) and genome build GRCh37/hg19. Pre-phasing and imputation were done using SHAPEIT and Eagle (v2.4), and variants with an R-square less than 0.20 were excluded.

Statistical Methods

The endpoint used for this GWA analysis was the first measured letrozole concentration, which was square-root transformed prior to analysis to improve data normality. Letrozole concentrations below the LLOQ (3 of 228 patients or 1.3%) were replaced with the LLOQ value (7.0 ng/mL). Each genotyped or imputed SNP was independently tested for association with letrozole concentration based on a genome-wide significance level of $5x10^{-8}$. All associations were analyzed under an additive genetic model using genotyped or imputed allelic dosages in PLINK 1.9 or 2.0[15], respectively. Associations were adjusted for BMI and age due to their previously reported effects on letrozole concentrations[10]. Additional GWA analyses were carried out after conditioning on the only significantly associated genotyped polymorphism in the primary GWA analysis (rs7937) to identify other genomic regions associated with letrozole concentrations. As previously described[10], patients were classified as normal, intermediate, and slow metabolizers based on known *CYP2A6* genotypes, and the primary GWA analysis of age- and BMI-adjusted letrozole levels was repeated after adjusting for CYP2A6 metabolizer phenotype. dbSNP and LDlink[16] were used to annotate variants and examine patterns of linkage disequilibrium in genomic regions of interest. Genotyped and imputed variants were further investigated using publicly available *in silico* tools. The Genotype-Tissue Expression (GTEx) project provides access to a database of studies relating genetic variants with measured gene expression within various tissue types.[17] RegulomeDB[18] scores genetic variants from 0.0 to 1.0 with higher scores indicating increased likelihood to be a regulatory variant.[19] Unless specified otherwise, all other analyses were carried out using a combination of in-house R programs and shell scripts.

Results

ELPh Patients and Letrozole Concentrations

Of the 503 patients enrolled on the ELPh study, 228 self-reported white patients randomized or cross-over to the letrozole arm had measured letrozole and genome-wide genetic data (Figure 1). Demographic and clinical data including letrozole concentrations from these patients are reported in Table 1 [10]. The median letrozole plasma concentration was 88 ng/mL (interquartile range of 46 ng/mL).

GWAS Results

In our sample of 228 patients, one genotyped variant (rs7937) was significantly associated with letrozole concentration before ($p=1.61x10^{-9}$) and after adjustment for age and BMI (reference/effect alleles: T/C, beta-coefficient (β)= 1.19, standard error (SE)=0.18, $p=5.26x10^{-10}$). Median (interquartile range) letrozole concentration in patients with genotypes CC, CT, and TT were 107.9 (48.5), 90.5 (46.4), and 73.2 (33.4) ng/ml,

respectively (Figure 2). This variant explained approximately 13% of the inter-individual variability in unadjusted letrozole plasma concentrations.

A list of all genotyped or imputed variants with $p<1x10^{-6}$ can be found in Supplementary Table 1 (Figure 3). Ten imputed variants in the chromosome 19 genomic region surrounding rs7937 exceeded genome-wide significance ($p<5x10^{-8}$) with age and BMI adjusted letrozole concentration (Table 2, Figure 4). All eleven variants, which were within 81 kb of each other, were well imputed (average imputation R-square of 0.885 to 0.999) and correlated with each other (linkage disequilibrium R^2 of 0.151 to 0.995). The most strongly associated imputed SNP was rs56113850 (reference/effect alleles: C/T, β=1.45, SE=0.19, $p=2.46x10^{-13}$, purple diamond in Figure 4) an intronic variant in moderate linkage ($r^2=0.21$) with rs7937.

None of these variants have obvious effects on CYP2A6 protein sequence, such as non-synonymous changes or introduction of a stop codon, nor are they correlated with known exonic functional *CYP2A6* SNPs. This suggests they may directly or indirectly affect the regulation or expression of CYP2A6. The genotyped variant, rs7937, is not associated with CYP2A6 expression in GTEx and has the lowest possible score (0.00, Table 2) for likelihood to be a regulatory variant in RegulomeDB. The imputed variant with the strongest association with letrozole concentration, rs56113850, has one of the highest RegulomeDB scores (0.60906) and is associated with *CYP2A6* expression in multiple tissues including liver in GTEx ($p=2.5x10^{-6}$, Table 2). The imputed variant with the highest RegulomeDB score was rs12461383 (0.70496), which was also associated with *CYP2A6* expression in liver tissue ($p=2.3x10^{-5}$) in GTEx.

Adjusting for *CYP2A6*-predicted metabolizer group (normal, intermediate, and slow), the association with age- and BMI-adjusted letrozole concentrations was nominally improved for rs7937 ($p=3.86x10^{-10}$) and slightly reduced below genome-wide significance for rs56113850 ($p=2.34x10^{-7}$). No new genome-wide significant signals emerged after conditioning on rs7937 or rs56113850 genotype. Finally, a variant previously implicated in anastrozole concentrations near the *SLC83A7* gene (rs11648166), [12] which was not genotyped but was well imputed (average imputation R-square 0.974), was not associated with letrozole concentration (p=0.34).

Discussion

Variability in systemic letrozole concentrations during treatment can be partially explained by inherited variation in genes relevant to the metabolism and transport of letrozole, including *CYP2A6*. Using a genome-wide approach, we confirmed that the *CYP2A6* region is the primary genetic determinant of letrozole concentrations in post-menopausal patients with HR+ breast cancer. Interestingly, our top hits are in non-exonic variants that are not typically included in CYP2A6 metabolic activity prediction[20] and may be associated with CYP2A6 expression and activity. We did not identify any additional genes that affect letrozole concentrations.

Our GWA results confirm findings from our prior candidate-gene study in the ELPh cohort, in which we reported that *CYP2A6* genetics was a critical determinant of letrozole concentrations[10]. In that study, we genotyped known functional variants in *CYP2A6* to predict each patient's metabolic phenotype, which explained approximately 20% of the variability in letrozole plasma concentrations. The non-exonic variants identified in the current GWAS, including rs7937 and rs56113850, explained approximately 15% of the overall variability in transformed and unadjusted letrozole concentrations. A combined model including CYP2A6 metabolic phenotype and either of our non-exonic variants explained nearly 40% of the variability in transformed letrozole concentrations, suggesting that the association of our non-exonic variants are completely independent of the variants included in our CYP2A6 metabolic phenotype prediction.

SNP rs7937, which is located on chromosome 19 approximately 50kb downstream of *CYP2A6*, has previously been identified in GWAS of CYP2A6-related phenotypes, including nicotine metabolism[21], and downstream clinical phenotypes, including chronic obstructive pulmonary disease[22] and lung cancer[21]. Functional studies indicate that this SNP affects methylation and expression of several genes in this region, such as *EGLN2*, but not *CYP2A6* [23], consistent with our in silico analysis of GTEx. Our search for additional variants within and around CYP2A6 identified several imputed SNPs with similar or stronger apparent associations. The imputed variant with the strongest association, rs56113850, is an intronic variant that has also been identified in several GWAS of smoking and nicotine metabolism[24-28] and is associated with *CYP2A6* gene expression in GTEx. Tanner et al. reported that the C allele is part of a haplotype with higher CYP2A6 protein expression and nicotine metabolic activity, indicating lower CYP2A6 expression and activity for the T allele.[25] Consistent with the results of Tanner et al., our study indicates lower metabolic activity for the T allele, as indicated by higher letrozole concentrations (β=1.45). Another imputed variant of potential interest is rs12461383, which has also been identified in GWAS of smoking-related phenotypes.[29] This variant is located approximately 10 KB upstream of *CYP2A6* in a region that has regulatory consequence in CHiP seq and DNase seq data within RegulomeDB, and was also associated with CYP2A6 gene expression in GTEx.[17,19] While our *in silico* analysis suggests rs56113850 may affect CYP2A6 activity and be responsible for these genetic associations, further functional studies are needed to confirm the functional variant.

Our findings indicate that any strategy to dose letrozole based on *CYP2A6* genotype should include these non-exonic variants in combination with variants typically considered in CYP2A6 metabolic activity phenotype prediction[20]. However, steady-state concentrations of letrozole, or any other AI, are not confirmed to have clinically relevant effects on treatment outcomes[6,8]. Prior work in the ELPh cohort was unable to demonstrate that patients with low AI concentrations had inferior estrogenic response[30] or that patients with high AI concentrations had worse toxicity during treatment[7]. Thus, there is currently no evidence that letrozole pharmacokinetics affects treatment outcomes and no rationale for testing CYP2A6 genotype or activity to inform letrozole dosing. If future analyses of large patient cohorts reveal that letrozole concentrations have a meaningful effect on treatment efficacy and/or toxicity, CYP2A6-guided dosing may be a useful strategy for personalized letrozole dosing to improve therapeutic outcomes in patients with HR+ breast cancer.

The strengths of this study included the use of a large cohort of prospectively accrued patients with standardized sample collection and analysis of letrozole plasma concentrations. Additionally, our use of an unbiased genome-wide approach and rigorous statistical adjustment confirm that variation in the *CYP2A6* region is the primary genetic driver of letrozole concentrations. However, our analysis has some limitations. Although this cohort is large for a pharmacokinetic study, it is relatively small compared to other GWAS, which results in a loss of power to detect genetic factors with smaller effect sizes. Additionally, our GWAS analyses was largely limited to common genetic variants (minor allele frequencies of at least 1%) that were genotyped or reliably imputed. Thus, we did not evaluate the role of rare variants in circulating letrozole concentrations. Additionally, the dosing and time of sample collection were not mandated by the protocol, so there is likely to be some pharmacokinetic variability caused by noncompliance with letrozole treatment or the recommended two-hour window between dosing and sample collection. Finally, we were not able to obtain an independent cohort of patients with measured letrozole concentrations to attempt validation of this pharmacogenetic association or conduct functional studies to determine which variant affects CYP2A6 expression or activity.

In conclusion, this GWAS confirms that the primary genetic driver of steady-state letrozole concentration in patients with HR+ breast cancer is in the *CYP2A6* region and suggests that non-exonic variants including rs56113850 may be markers for CYP2A6 expression and activity. These variants should be considered for inclusion in systems that translate *CYP2A6* genetics to metabolic activity phenotype.[20] If future studies demonstrate that circulating letrozole concentrations affect efficacy or toxicity of treatment, *CYP2A6* genetics may be useful to individualize letrozole dosing to improve clinical outcomes in patients with HR+ breast cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported by Pharmacogenetics Research Network Grant No. U-01 GM61373 and Clinical Pharmacology Training Grant No. 5T32-GM08425 (both awarded to David.A.Flockhart) from the National Institute of General Medical Sciences, National Institutes of Health (NIH), from Grants No. M01-RR000042 (University of Michigan), M01-RR00750 (Indiana University), and M01-RR00052 (Johns Hopkins University) from the National Center for Research Resources (NCRR), a component of the NIH, the Breast Cancer Research Foundation (BCRF) (N003173 to JMR and DFH), the National Cancer Institute (5T32CA083654), the National Institute of General Medical Sciences (GM099143 to J.M.R.) and the National Institutes of Health through the University of Michigan's Cancer Center Support Grant (P30 CA046592) by the use of the following Cancer Center Core: University of Michigan DNA Sequencing Core. In addition, these studies were supported by grants from Pfizer (D.F.H.), Novartis Pharma AG (D.F.H.), the Fashion Footwear Association of New York/QVC Presents Shoes on Sale (D.F.H.). Drugs were supplied by Novartis and Pfizer. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 07/23/2020.

References:

1. Goss PE, Ingle JN, Martino S, Robert NJ, Muss HB, Piccart MJ, et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med. 2003; 349 (19):1793–1802. [PubMed: 14551341]
2. Early Breast Cancer Trialists' Collaborative G, Dowsett M, Forbes JF, Bradley R, Ingle J, Aihara T, et al. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet (London, England). 2015; 386 (10001):1341–1352.
3. Geisler J, Haynes B, Anker G, Dowsett M, Lonning PE. Influence of letrozole and anastrozole on total body aromatization and plasma estrogen levels in postmenopausal breast cancer patients evaluated in a randomized, cross-over study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2002; 20 (3):751–757. [PubMed: 11821457]
4. Henry NL, Giles JT, Stearns V. Aromatase inhibitor-associated musculoskeletal symptoms: etiology and strategies for management. Oncology (Williston Park, NY). 2008; 22 (12):1401–1408.
5. Henry NL, Azzouz F, Desta Z, Li L, Nguyen AT, Lemler S, et al. Predictors of aromatase inhibitor discontinuation as a result of treatment-emergent symptoms in early-stage breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2012; 30 (9):936–942. [PubMed: 22331951]
6. Ingle JN, Kalari KR, Buzdar AU, Robson ME, Goetz MP, Desta Z, et al. Estrogens and their precursors in postmenopausal women with early breast cancer receiving anastrozole. Steroids. 2015; 99 (Pt A):32–38. [PubMed: 25163006]
7. Kadakia KC, Snyder CF, Kidwell KM, Seewald NJ, Flockhart DA, Skaar TC, et al. Patient-Reported Outcomes and Early Discontinuation in Aromatase Inhibitor-Treated Postmenopausal Women With Early Stage Breast Cancer. Oncologist. 2016; 21 (5):539–546. [PubMed: 27009936]
8. Hertz DL, Henry NL, Rae JM. Germline genetic predictors of aromatase inhibitor concentrations, estrogen suppression and drug efficacy and toxicity in breast cancer patients. Pharmacogenomics. 2017; 18 (5):481–499. [PubMed: 28346074]
9. Hertz DL, Kidwell KM, Seewald NJ, Gersch CL, Desta Z, Flockhart DA, et al. Polymorphisms in drug-metabolizing enzymes and steady-state exemestane concentration in postmenopausal patients with breast cancer. Pharmacogenomics J. 2017; 17 (6):521–527. [PubMed: 27549341]
10. Desta Z, Kreutz Y, Nguyen AT, Li L, Skaar T, Kamdem LK, et al. Plasma letrozole concentrations in postmenopausal women with breast cancer are associated with CYP2A6 genetic variants, body mass index, and age. Clinical pharmacology and therapeutics. 2011; 90 (5):693–700. [PubMed: 21975350]
11. Motsinger-Reif AA, Jorgenson E, Relling MV, Kroetz DL, Weinshilboum R, Cox NJ, et al. Genome-wide association studies in pharmacogenomics: successes and lessons. Pharmacogenet Genomics. 2013; 23 (8):383–394. [PubMed: 20639796]
12. Dudenkov TM, Liu D, Cairns J, Devarajan S, Zhuang Y, Ingle JN, et al. Anastrozole Aromatase Inhibitor Plasma Drug Concentration Genome-Wide Association Study: Functional Epistatic Interaction Between SLC38A7 and ALPPL2. Clin Pharmacol Ther. 2019; 106 (1):219–227. [PubMed: 30648747]
13. Henry NL, Jacobson JA, Banerjee M, Hayden J, Smerage JB, Van Poznak C, et al. A prospective study of aromatase inhibitor-associated musculoskeletal symptoms and abnormalities on serial high-resolution wrist ultrasonography. Cancer. 2010; 116 (18):4360–4367. [PubMed: 20549827]
14. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet (Research Support, N.I.H., Extramural Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't). 2016; 48 (10):1284–1287. doi: 1210.1038/ng.3656. Epub 2016 Aug 1229. [PubMed: 27571263]
15. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience (Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't). 2015; 4:7. (doi):10.1186/s13742-13015-10047-13748. eCollection 12015. [PubMed: 25722852]
16. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics

(Research Support, N.I.H., Intramural). 2015; 31 (21):3555–3557. doi: 3510.1093/bioinformatics/btv3402. Epub 2015 Jul 3552. [PubMed: 26139635]

17. The Genotype-Tissue Expression (GTEx) project. Nat Genet (Research Support, N.I.H., Extramural Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't). 2013; 45 (6):580–585. doi: 510.1038/ng.2653. [PubMed: 23715323]
18. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res (Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't). 2012; 22 (9):1790–1797. doi: 1710.1101/gr.137323.137112. [PubMed: 22955989]
19. Dong S, Boyle AP. Predicting functional variants in enhancer and promoter elements using RegulomeDB. Hum Mutat (Research Support, N.I.H., Extramural). 2019; 40 (9):1292–1298. doi: 1210.1002/humu.23791. Epub 22019 Jun 23722. [PubMed: 31228310]
20. Lee S-B, Wheeler MM, Thummel KE, Nickerson DA. Calling Star Alleles With Stargazer in 28 Pharmacogenes With Whole Genome Sequences. Clinical pharmacology and therapeutics. 2019; 106 (6):1328–1337. [PubMed: 31206625]
21. Timofeeva MN, McKay JD, Smith GD, Johansson M, Byrnes GB, Chabrier A, et al. Genetic polymorphisms in 15q25 and 19q13 loci, cotinine levels, and risk of lung cancer in EPIC. Cancer Epidemiol Biomarkers Prev. 2011; 20 (10):2250–2261. [PubMed: 21862624]
22. Cho MH, Castaldi PJ, Wan ES, Siedlinski M, Hersh CP, Demeo DL, et al. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. Hum Mol Genet. 2012; 21 (4):947–957. [PubMed: 22080838]
23. Nedeljkovic I, Lahousse L, Carnero-Montoro E, Faiz A, Vonk JM, de Jong K, et al. COPD GWAS variant at 19q13.2 in relation with DNA methylation and gene expression. Hum Mol Genet. 2018; 27 (2):396–405. [PubMed: 29092026]
24. Chen LS, Hartz SM, Baker TB, Ma Y, N LS, Bierut LJ. Use of polygenic risk scores of nicotine metabolism in predicting smoking behaviors. Pharmacogenomics (Meta-Analysis Research Support, N.I.H., Extramural). 2018; 19 (18):1383–1394. doi: 1310.2217/pgs-2018-0081. Epub 2018 Nov 1316. [PubMed: 30442082]
25. Tanner JA, Zhu AZ, Claw KG, Prasad B, Korchina V, Hu J, et al. Novel CYP2A6 diplotypes identified through next-generation sequencing are associated with in-vitro and in-vivo nicotine metabolism. Pharmacogenet Genomics. 2018; 28 (1):7–16. doi: 10.1097/FPC.0000000000000317. [PubMed: 29232328]
26. Patel YM, Park SL, Han Y, Wilkens LR, Bickeböller H, Rosenberger A, et al. Novel Association of Genetic Markers Affecting CYP2A6 Activity and Lung Cancer Risk. Cancer Res (Research Support, N.I.H., Extramural Research Support, N.I.H., Intramural Research Support, Non-U.S. Gov't). 2016; 76 (19):5768–5776. doi: 5710.1158/0008-5472.CAN-5716-0446. Epub 2016 Aug 5763. [PubMed: 27488534]
27. Baurley JW, Edlund CK, Pardamean CI, Conti DV, Krasnow R, Javitz HS, et al. Genome-Wide Association of the Laboratory-Based Nicotine Metabolite Ratio in Three Ancestries. Nicotine Tob Res (Twin Study). 2016; 18 (9):1837–1844. doi: 1810.1093/ntr/ntw1117. Epub 2016 Apr 1825.
28. Loukola A, Buchwald J, Gupta R, Palviainen T, Hällfors J, Tikkanen E, et al. A Genome-Wide Association Study of a Biomarker of Nicotine Metabolism. PLoS Genet (Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't). 2015; 11 (9):e1005498. doi: 1005410.1001371/journal.pgen.1005498. eCollection 1002015. [PubMed: 26407342]
29. Lutz SM, Cho MH, Young K, Hersh CP, Castaldi PJ, McDonald ML, et al. A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. BMC Genet (Clinical Trial Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't). 2015; 16:138. (doi):10.1186/s12863-12015-10299-12864. [PubMed: 26634245]
30. Hertz DL, Speth KA, Kidwell KM, Gersch CL, Desta Z, Storniolo AM, et al. Variable aromatase inhibitor plasma concentrations do not correlate with circulating estrogen concentrations in post-menopausal breast cancer patients. Breast Cancer Res Treat. 2017; 165 (3):659–668. [PubMed: 28643023]

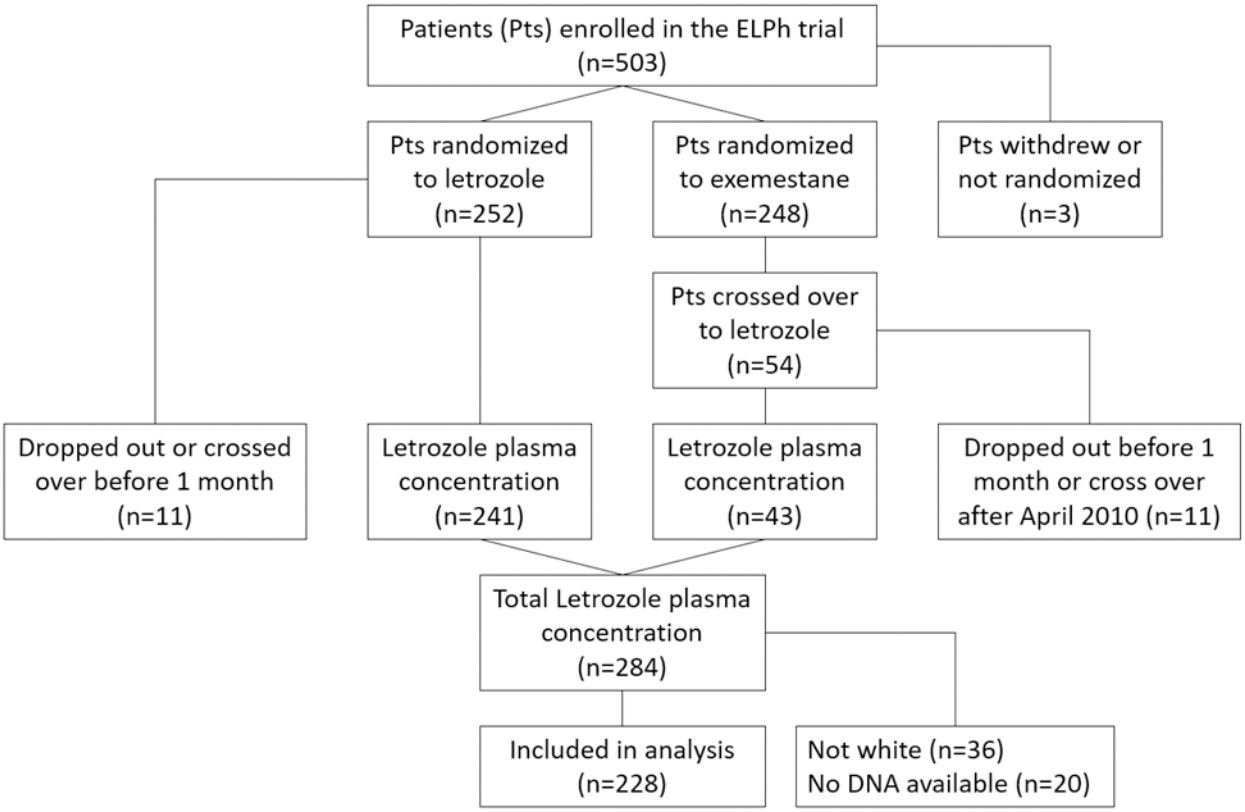

Figure 1:
CONSORT Diagram Illustrating Patient Matriculation from the ELPh study to this GWAS

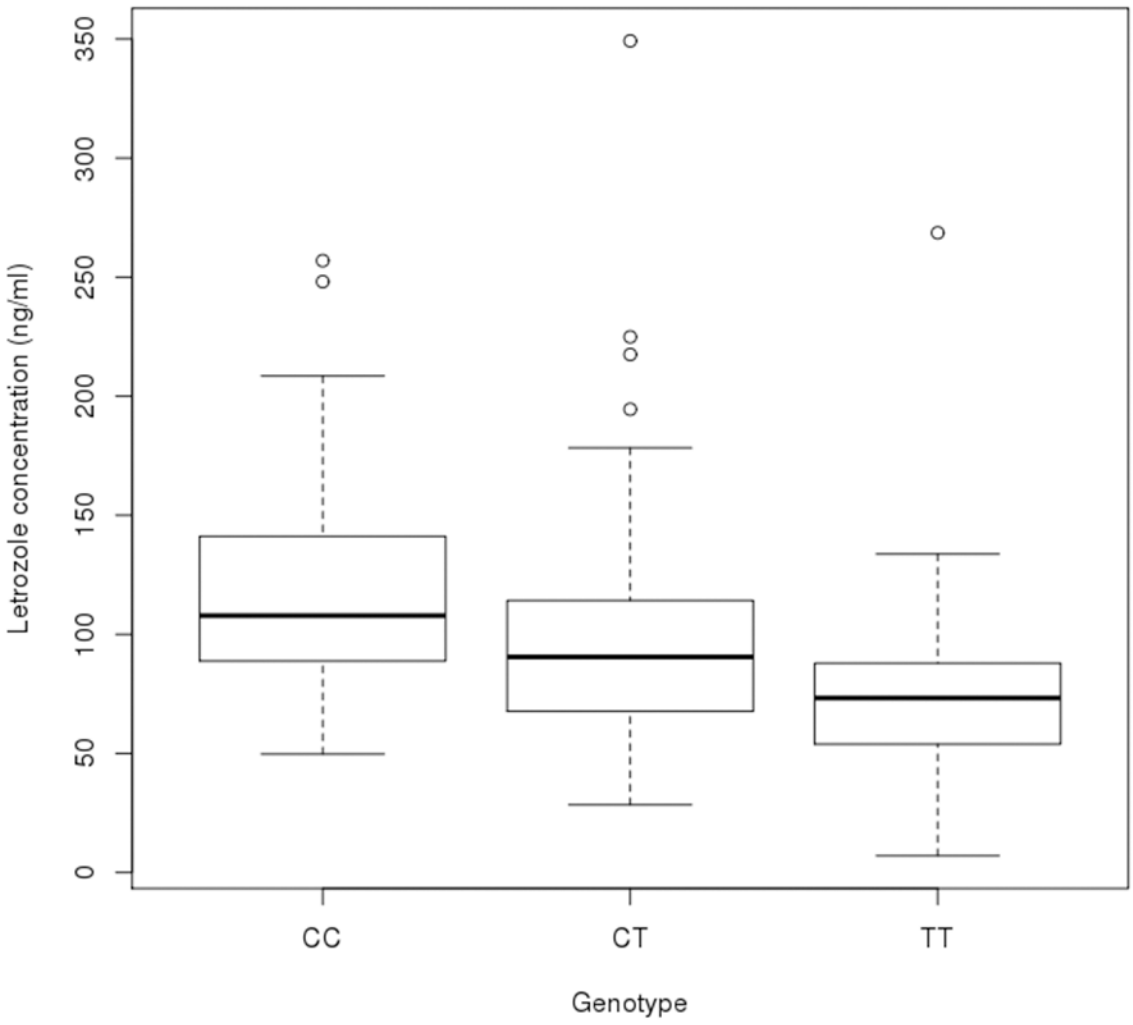

Figure 2:

Letrozole concentrations stratified by rs7937 genotype. Letrozole concentrations were higher in carriers of the rs7937 effect (C) allele (additive $p=6.79x10^{-10}$). Median (n, ± interquartile range) letrozole concentration in patients with genotypes CC, CT, and TT were 107.9 (n=48, ± 48.5), 90.5 (n=118, ± 46.4), and 73.2 (n=, 62 ±33.4) ng/ml, respectively.

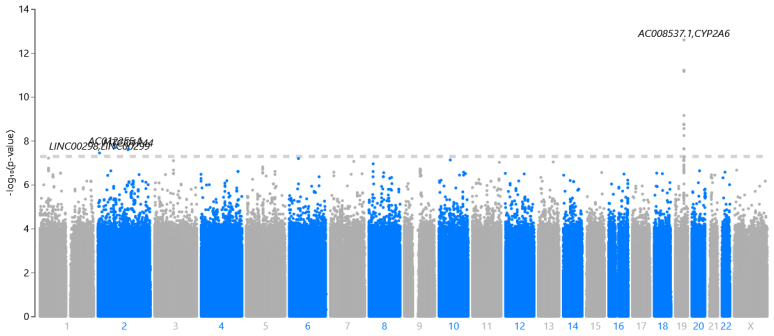

Figure 3.

Association between letrozole concentrations and variants that were genotyped or imputed. The association for one genotyped single nucleotide polymorphisms (SNP), rs7937, with steady-state letrozole concentrations ($P=1.61\times10^{-9}$) surpassed the genome-wide significance threshold of 5×10^{-8} indicated by the horizontal line. Including the imputed variants, 15 SNPs surpassed genome-wide significance, including 11 in the genomic region around rs7937.

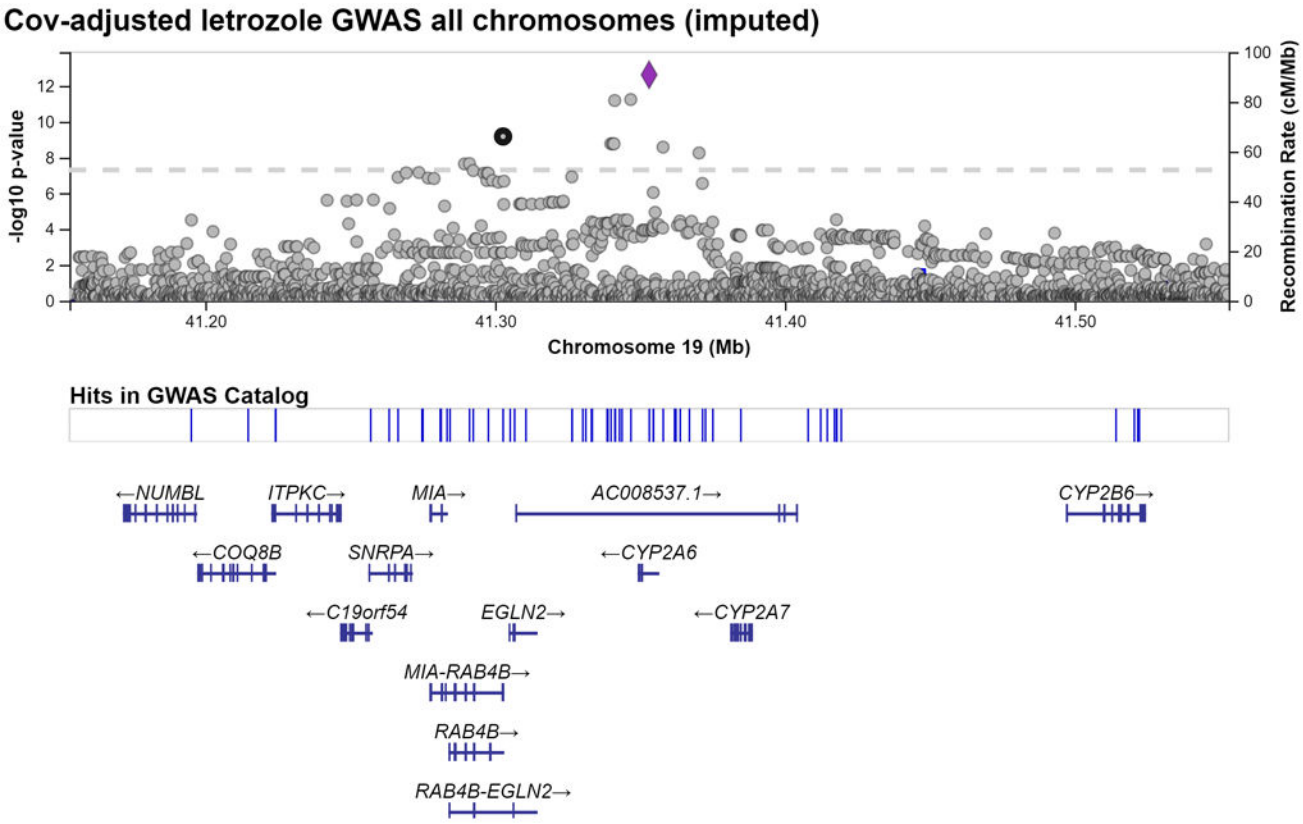

Figure 4:

Locus zoom plot around the most strongly associated variant, rs56113850, p=2.46x10^{-13}, purple diamond). The only SNP genotyped in this region, rs7937, is indicated with the bolded circle.

Table 1.

Characteristics of breast cancer patients included in this analysis (n=228)

Trait	N (%) or Median (IQR[1])
Age at enrollment (years)	59.5 (33)
Body mass index (kg/m^2)	28.5 (9.1)
Letrozole level (ng/ml)	87.7 (45.8)
Letrozole measured at 3 months	205 (90%)
Prior chemotherapy treatment	108 (47%)

IQR: Interquartile range.

Table 2.

Imputed variants in Chromosome 19 significantly associated with age- and BMI-adjusted letrozole plasma concentrations*

Variant	Position[a]	Nearest genes	Alleles[b]	EAF[c]	r^2	Beta	SE	P-value	RegulomeDB[d]	GTEx[e]
rs12973666	41289397	*RAB4B-EGLN2* (intron)	C/G	0.499	0.81	1.06	0.18	$2.30x10^{-08}$	0.55436	--
rs17726276	41291119	*RAB4B-EGLN2* (intron)	A/G	0.500	0.82	1.06	0.18	$2.20x10^{-08}$	0.55436	--
rs7937	41302706	*RAB4B-EGLN2* (intron)	T/C	0.464	1.00	1.19	0.18	$6.79x10^{-10}$	0.0	--
rs12459249	41339896	Intergenic	C/T	0.389	0.20	1.29	0.21	$1.72x10^{-09}$	0.60906	--
rs10853742	41340573	Intergenic	C/G	0.393	0.20	1.29	0.21	$1.75x10^{-09}$	0.60906	--
rs11667314	41340983	Intergenic	C/T	0.392	0.20	1.29	0.21	$1.74x10^{-09}$	0.50689	--
rs12461964	41341229	Intergenic	G/A	0.542	0.33	1.34	0.19	$6.70x10^{-12}$	0.005	--
rs2316205	41346768	Intergenic	C/T	0.543	0.31	1.35	0.19	$5.89x10^{-12}$	0.14	--
rs56113850	41353107	*CYP2A6* (intron)	C/T	0.460	0.21	1.45	0.19	$2.46x10^{-13}$	0.60906	$2.5x10^{-6}$
rs57837628	41357910	*CYP2A6* (2 KB upstream)	G/A	0.506	0.24	1.17	0.19	$2.66x10^{-09}$	0.60906	$1.9x10^{-5}$
rs12461383	41370338	*CYP2A6* (10 KB upstream)	G/C	0.501	0.24	1.14	0.19	$5.63x10^{-09}$	0.70496	$2.3x10^{-5}$

*Square root-transformed letrozole concentrations.

Acronyms: EAF: Effect allele frequency. SE: Standard error

[a]Position based on genome build 37.

[b]Effect allele is second allele.

[c]EAF in the Haplotype Reference Consortium European reference panel.

[d]Score from RegulomeDB on a scale of 0.0-1.0 with higher numbers indicating higher likelihood to be a regulatory variant.

[e]Association of variant with CYP2A6 expression in liver. – Indicates no association with CYP2A6 in any tissue